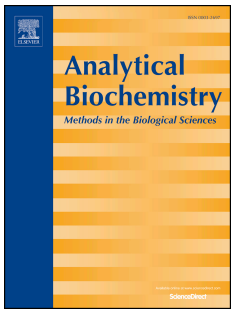
Accepted Manuscript



A recombinant fusion protein-based, fluorescent protease assay for high throughput-compatible substrate screening

Beáta Bozóki, Lívia Gazda, Ferenc Tóth, Márió Miczi, János András Mótyán, József Tőzsér

PII: S0003-2697(17)30430-X

DOI: 10.1016/j.ab.2017.11.001

Reference: YABIO 12832

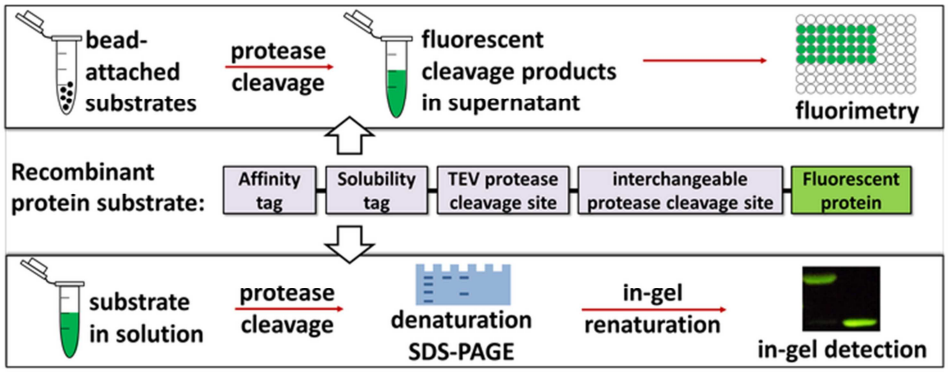
To appear in: *Analytical Biochemistry*

Received Date: 10 July 2017

Revised Date: 2 November 2017

Accepted Date: 4 November 2017

Please cite this article as: Beá. Bozóki, Lí. Gazda, F. Tóth, Máó. Miczi, Já.Andrá. Mótyán, Jó. Tőzsér, A recombinant fusion protein-based, fluorescent protease assay for high throughput-compatible substrate screening, *Analytical Biochemistry* (2017), doi: 10.1016/j.ab.2017.11.001.



A recombinant fusion protein-based, fluorescent protease assay for high throughput-compatible substrate screening

Beáta Bozóki[1,2], Lívia Gazda[1], Ferenc Tóth[1], Márió Miczi[1], János András Mótyán[1] and József Tőzsér[1,*]

[1]*Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Debrecen, Hungary*

[2]*Biotechnological Analytical Department, Gedeon Richter Plc, 19-21. Gyömrői Rd. Budapest, H-1103 Hungary.*

***Corresponding author:**

Dr. József Tőzsér,

Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Debrecen, Hungary; H-4012 Debrecen, POB 6. Debrecen, Hungary. Tel.: +36-52/416-432; Fax: +36-52/314989.

Email: tozser@med.unideb.hu

Subject category: Enzymatic assays and analysis

Contact information of authors:

Beáta Bozóki: bobea86@gmail.com

Lívia Gazda: gazda.livia@med.unideb.hu

Ferenc Tóth: tferenc@med.unideb.hu

Márió Miczi: miczimario@med.unideb.hu

János András Mótyán: motyan.janos@med.unideb.hu

József Tőzsér: tozser@med.unideb.hu

Abstract

In connection with the intensive investigation of proteases, several methods have been developed for analysis of the substrate specificity. Due to the great number of proteases and the expected target molecules to be analyzed, time- and cost-efficient high-throughput screening (HTS) methods are preferred. Here we describe the development and application of a separation-based HTS-compatible fluorescent protease assay, which is based on the use of recombinant fusion proteins as substrates of proteases.

The protein substrates used in this assay consists of N-terminal (hexahistidine and maltose binding protein) fusion tags, cleavage sequences of the tobacco etch virus (TEV) and HIV-1 proteases, and a C-terminal fluorescent protein (mApple or mTurquoise2). The assay is based on the fluorimetric detection of the fluorescent proteins, which are released from the magnetic bead-attached substrates by the proteolytic cleavage. The protease assay has been applied for activity measurements of TEV and HIV-1 proteases to test the suitability of the system for enzyme kinetic measurements, inhibition studies, and determination of pH optimum. We also found that denatured fluorescent proteins can be renatured after SDS-PAGE of denaturing conditions, but showed differences in their renaturation abilities. After in-gel renaturation both substrates and cleavage products can be identified by in-gel UV detection.

Keywords:

recombinant fusion protein substrate;

protease assay;

HTS-compatible;

fluorescent protein;

human immunodeficiency virus type 1 protease;

tobacco etch virus protease

1. Introduction

Due to their essential role in numerous biological processes, proteases are claimed to be one of the most important groups among known enzyme families. Proteases are involved in many biological pathways; their malfunctions can be responsible for many degenerative disorders, cancer, inflammations, etc. On the other hand, proteases are important drug targets due to their central role in the life-cycle of different bacteria, viruses, and other parasites. In humans, 5-10% of drug targets show protease activity, and the search for new molecules to target proteases is still highly important both in academic and industrial research (**Choe et al., 2006; Kostallas et al., 2011; Zhang et al., 2012**).

In connection with the intensive investigation of proteases, several methods have been developed for the analysis of their substrate specificity. Due to the great number of the proteases and the expected target molecules to be analyzed, time- and cost efficient high-throughput screening (HTS) methods are preferred. According to Zhang (**2012**), HTS-protease assays can be divided into two major groups: homogenous and separation-based assays.

In the case of homogenous assays, it is not required for the detection to separate the products and/or the substrates present in the reaction mixture. Fluorescent detection-based assays are highly preferred due to their high sensitivity that allows detection in a dense, low volume format, which is especially important in HTS. The most widely utilized fluorescence-based approaches include assays based on (*i*) chemically quenched dyes for example in the assay developed by Hassiepen et al. (**2007**) for deubiquitinating proteases; (*ii*) fluorescence resonance energy transfer (FRET), of which the principles were firstly described by Matayoshi et al. (**1990**) on HIV PR and it has become widely utilized and also further improved for example by Branchini et al. (**2011**); (*iii*) dual-label quenched pairs such as the one developed in our laboratory by Bagossi et al. (**2004**); or (*iv*) fluorescence polarization for example in case of the assay designed by Levine et al. (**1997**) for human cytomegalovirus protease.

On the other hand, separation-based assays offer another excellent way for investigating protease activity. In these approaches products and/or substrates are isolated from the reaction mixture mainly by liquid chromatography and analyzed independently (e.g.: by mass spectrometry). This assay type utilizing reversed-phase high-performance liquid chromatography (RP-HPLC) is also commonly applied in our laboratory as described by

Tözsér et al. (**1991**) or Fehér et al. (**2006**). A great advantage of these assay types is that they allow kinetic reading of the enzymatic reactions, although instrumentation requirements of these approaches are relatively high.

All of these approaches are excellently engineered and can be well-applied in several cases, although each of them has its own limitation. Chemically quenched dye-based assays can monitor the changes in substrate sequence only on the amino (N) side of the scissile bond, and cannot be applied for proteases that have strict requirements in the carboxyl (C) side of the cleavage site. FRET-based approaches require large percent of substrate turnover for reliable readouts, while assays using dual-label quenched pairs often limited by high background signals due to the improper separation of the fluorescents spectra. The fluorescence polarization assays have high sensitivity for the changes of the fluorescent signal, however, outside a moderate range the signal does not correlate linearly with product concentration. Moreover, in several assay formats the possibility of getting false positive results is relatively high, due to compound aggregation, enzyme degradation, impurities, or chemical reactivity, especially in the case of cysteine proteases, where thiol/thiolate group causes high-reactivity (**Jadhav et al., 2010; Zhang et al., 2012**).

In addition to the widely utilized short peptide-based substrates or small molecule fluorescent probes, recombinant fusion protein substrates have been also successfully applied in several protease assay formats (**Patel et al., 2001; Chaparro-Riggers et al., 2005; Askin et al., 2011; Branchini et al., 2011; Zhou et al., 2014**). Moreover, fluorescent protein (FP) containing fusion protein substrates offer highly sensitive, cost-efficient, and stable tools, which can be easily adapted for the detection of site-specific protease cleavage (**Patel et al., 2001; Chaparro-Riggers et al., 2005; Aoki et al., 2008; Askin et al., 2011; Branchini et al., 2011**). Some of these assays are also compatible with HTS formats (**Chaparro-Riggers et al., 2005; Askin et al., 2011)**, and some also allow the determination of cleavage kinetics (**Chaparro-Riggers et al., 2005**).

All of these approaches comprise excellent solutions for studying protease specificity, however, considering all of the listed aspects the development of assay formats that integrate most of the advantages of these platforms and may offer new ones in an improved format is still on demand.

Here we report the development and application of a separation-based, HTS-compatible, *in vitro* fluorescent proteolytic assay system, which is based on the use of recombinant fusion protein substrates. Substrates tagged with mApple or mTurquoise2

fluorescent proteins have been tested for activity measurements of the human immunodeficiency virus type 1 protease (HIV-1 PR) and the tobacco etch virus protease (TEV PR). The designed system was applied to determine kinetic parameters of the studied enzymes, and to test the inhibitory effect of the protease inhibitor amprenavir on HIV-1 PR and the pH dependence of TEV PR activity. Besides measuring the fluorescence of the supernatants (due to the release of fluorescent cleavage products from the magnetic bead-attached substrates) by a fluorimeter, the substrates and cleavage products were also analyzed by polyacrylamide gel electrophoresis in the presence or absence of denaturing conditions. We found that removal of SDS from the gel enables the system for fluorescent in-gel detection of the proteins together with their molecular weight-based identification. The recombinant protein substrate system described here is suitable to generate substrate libraries or insert wide variety of sequences into the substrates to study several types of proteases.

2. Materials and methods

All materials were purchased from Sigma-Aldrich (St Louis, MO, USA) unless otherwise indicated.

Oligonucleotide primer sequences are also available in the public oligonucleotide database of Laboratory of Retroviral Biochemistry (http://lrb.med.unideb.hu/research/oligos).

2.1 Expression vectors

pDest-His_6-MBP-mApple and pDest-His_6-MBP-mTurquoise2 plasmids were prepared by Gateway Cloning Technology (Thermo Fischer Scientific, Invitrogen) based on Tropea et al. (**2007**). The linear DNA sequences to be transferred via pDON221 donor vector into pDEST-His_6-MBP (**Tropea et al., 2007**) were amplified by a two-step PCR reaction. pDON221 and pDEST-His_6-MBP vectors were kind gifts of dr. David S. Waugh (NCI-Frederick, USA).

In the first step the coding sequences of the desired fluorescent proteins were amplified by a PCR reaction (100 µl final volume) using 300-300 ng N1 and C primers (5'-GAGAACCTGTACTTCCAGGGTGGTGCATCTACACGCGGTTTAATTAATCTTCTGCTAGCATGGTGAGCAAGGGCG-3') and C primers (5'-GGGGACACACTTTGTACAAGAAAGCTGGGTTATTACTTGTACAGCTCGTCCAT-3'), and 1800 ng mApple-N1 or pPalmitoyl-mTurquoise2 plasmids (Addgene) as templates. Reaction mixtures contained Pfu buffer + $MgSO_4$ (10x) (Thermo Fischer Scientific, Fermentas), Pfu DNA polymerase (Thermo Fischer Scientific, Fermentas), 10 mM dNTP mix, and nuclease-free water (to set final volume to 100 µl). The following PCR protocol was used: cycle1 (1x): 5 min at 94°C; cycle2 (30x): 1 min at 94°C, 1 min at 45°C, and 3 min at 72°C; cycle3 (1x): store at 4°C.

The PCR products were separated by gel electrophoresis using 1% agarose gel. The linear PCR products (having ~800 base pair lengths) were cut out from the gel and purified by Qiagen DNA extraction kit (Qiagen). Twenty-five ng of the purified DNA was used as a template in the second PCR reaction (100 µl final volume) containing 300 ng N2 (Thermo Fischer Scientific, Invitrogen; 5'-GGGGACAAGTTTGTACAAAGCAGGCTCGGAGAACCTGTACTTCCAG-3') and C primers, and Pfu buffer + $MgSO_4$ (10x) (Thermo Fischer Scientific, Fermentas), Pfu DNA polymerase (Thermo Fischer Scientific, Fermentas), 10 mM dNTP (Biorad) mix, and

nuclease-free water (to set final volume to 100 µl). The PCR protocol was the same as in the case of the first-step reaction.

The PCR products were purified by Qiagen Nucleotide Removal Kit (Qiagen). Three-hundred ng of the purified linear DNA sequence flanked by *attB* sites was used in a BP reaction (recombination between *attB* and *attP* sites) with pDON221 (Thermo Fischer Scientific, Invitrogen). Thereafter, LR reaction (for the recombination between *attL* and *attR* sites) was performed for 4 hours using 300 ng of the purified entry clone and 800 ng pDEST-His_6-MBP destination vector (Addgene). Gateway cloning steps and the following transformations were carried out according to the manufacturer's protocol.

2.2 'Cloning cassette' and the inserted HIV-1 PR cleavage site

Circular pDest-His_6-MBP-mApple and pDest-His_6-MBP-mTurquoise2 expression vectors were cleaved by PacI and NheI restriction endonucleases (New England Biolabs). Cleavage products were separated by 1% agarose gel electrophoresis, and the linear plasmids were cut out from the gel and purified by Qiagen DNA extraction kit (Qiagen). The annealing of the complement oligonucleotides coding for the cleavage sites to be inserted, and their ligation into the linearized expression plasmids were carried out by the following reactions: 200-200 ng forward (5'-TAAAGTGAGCCAGAACTATCCGATTGTGCAGG-3') and reverse (5'-CTAGCCTGCACAATCGGATAGTTCTGGCTCACTTTAAT-3') primers were mixed with 150 ng purified linearized expression plasmid. The reaction mixtures were incubated at 65°C for 2 minutes and after that at 4°C for 2 minutes. Thereafter, T4 DNA Ligase Reaction Buffer (10x) and T4 DNA Ligase (New England Biolabs) were added, and the reaction mixture was incubated at 25°C for 1 hour followed by incubation at 16°C overnight.

Next day 100 µl of *Escherichia coli* BL21(DE3) competent cells (Thermo Fischer Scientific, Invitrogen) were transformed by 15 µl of the ligation reaction mixture by heat shock (at 42°C). The transformation reaction was spread on Luria-Bertani (LB) agar plates containing 100 µg/ml ampicillin. Grown colonies were cultured at 37°C for overnight in LB medium containing 100 µg/ml ampicillin followed by plasmid preparation using Qiaprep Spin Mini Prep Kit (Qiagen). DNA sequences of the purified plasmids were verified by capillary DNA sequencing using pDest-His6-MBP-FP sequencing forward primer (5'-GATGAAGCCCTGAAAGACGCGCAG-3').

2.3 Expression of substrates

Recombinant substrates were expressed in *Escherichia coli* BL21(DE3) cells. Cells were grown in LB medium containing 100 µg/ml ampicillin at 37°C up to an absorbance of 0.6-0.8 at 600 nm. Protein expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) followed by incubation for 3 hours. Cells were harvested by centrifugation at 4,000 g for 15 minutes at 4°C (Jouan CR 412). Cell pellets were stored at least for 1 hour at -70°C before lysis of the cells, and were let to thaw on ice for 15 minutes. Pellets were suspended in 2 ml lysis buffer (50 mM sodium-acetate, 300 mM NaCl, 10 mM imidazole, 0.05% Tween 20, pH 8.0) containing phenylmethanesulfonyl-fluoride (PMSF) at a final concentration of 25 µg/ml. Lysozyme and DNase enzymes (New England Biolabs) were added in 1 mg/ml and 10 U/ml final concentrations, respectively. Cell suspensions were vortexed and occasionally mixed during incubation on ice for 10 minutes, then were aliquoted into microcentrifuge tubes and sonicated for 3 minutes. Tubes were centrifuged at 10,000 g for 20 minutes (Eppendorf 5415D) at room temperature, and the supernatants (cleared bacterial lysate containing the desired recombinant fusion substrate) were collected.

2.4 Purification of the substrates

Cleared bacterial lysates containing the desired recombinant fusion substrate were added to Ni-NTA magnetic agarose beads (Qiagen), the suspensions were incubated at least for 20 minutes at room temperature while continuously shaking. Magnetic beads were washed three times with 1% Tween 20 (pH 7.0), then washed three times by washing buffer (50 mM sodium-acetate, 300 mM NaCl, 5 mM imidazole, 0.05% Tween 20, pH 7.0) and three times by cleavage buffer (50 mM sodium-acetate, 300 mM NaCl, 0.05% Tween 20, pH 7.0) using Dynamag^{TM}-2 magnetic particle concentrator (Thermo Fischer Scientific, Invitrogen). Beads were incubated in cleavage buffer for 60 minutes at 37°C while continuously shaking at 600 rpm and thereafter beads were applied to Dynamag^{TM}-2 magnetic particle concentrator (Thermo Fischer Scientific, Invitrogen). Supernatants were removed and substrates were eluted from the beads using elution buffer (100 mM EDTA, 0.05% Tween 20, pH 8.0). Buffer was exchanged to cleavage buffer by using 10K Amicon tubes (Merck-Millipore). Total protein concentrations were determined by BCA protein assay (Thermo Fischer Scientific, Pierce) and by measuring absorbance at 280 nm by NanoDrop 2000 equipment (Thermo Fischer Scientific). Both methods were applied simultaneously to verify the results, but OD 280 nm values were used for further calculations.

2.5 Expression and purification of HIV-1 PR

The plasmid coding for the HIV-1 PR was kindly provided by Dr. John M. Louis (Laboratory of Chemical Physics, NIDDK, NIH). The HIV-1 PR coding region contained five stabilizing mutations (Q7K, L33I, L63I, C67A, and C95A), the construction of the plasmid was described previously (**Louis et al., 1999; Mahalingam et al., 1999**). HIV-1 PR was expressed in *Escherichia coli* BL21(DE3) cells. Cells were grown at 37°C up to an absorbance of 0.6–0.8 at 600 nm, in LB medium containing 100 µg/ml ampicillin, and induced for expression with 1 mM IPTG for 3 hours. Cells were harvested by centrifugation at 6,000 g for 20 min at 4°C. After removal of the supernatant, the cell pellet was treated as described previously (**Mahalingam et al., 2001**). Pellet was suspended in 20 volume of buffer A (50 mM Tris, 10 mM EDTA, pH 8.2), and lysed in the presence of 100 µg/ml lysozyme by sonication on ice. The lysate was centrifuged at 20,000 g for 20 min at 4°C, suspended in buffer B (buffer A containing 2 M guanidine-HCl and 1% Triton X-100) and centrifuged again. Pelleted inclusion bodies were suspended in buffer A and centrifuged again. The final pellet was dissolved in buffer C (50 mM Tris, 5 mM EDTA, 7.5 M guanidine-HCl, pH 8.0), filtered through 0.22 µm pore size filter (Merck-Millipore) and applied to a Superose 12 10/300 GL column (GE Healthcare) equilibrated in 50 mM Tris, 4 M guanidine-HCl, 5 mM EDTA containing buffer (pH 8.0) at a flow rate of 0.5 ml/min at ambient temperature. Peak fractions were pooled and subjected to RP-HPLC on a POROS 20 R2 column (Thermo Fischer Scientific, Applied Biosystems). Purity of selected fractions was assessed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) using 16% polyacrylamide gels. The protein was folded by dialysis into 0.05 M formic acid at pH 2.8 followed by a dialysis using cleavage buffer (50 mM sodium-acetate, 300 mM NaCl, 0.05% Tween 20, pH 7.0) or 2x cleavage buffer for HPLC assays.

2.6 Gel electrophoresis

Purified recombinant protein substrates and cleavage products were analyzed by PAGE, using 14% polyacrylamide gel. Two different protocols have been used: i) 6X loading buffer (300 mM Tris, 20% glycerol, 0.05% bromophenol blue, pH 6.8) containing no reducing agents was added to those samples, which were not heat-treated before the electrophoresis (**Gross et al., 2000**); ii) the samples mixed with 6X loading buffer - containing 12% SDS and 100 mM β-mercaptoethanol (β-ME) as reducing agents - were

heated at 95°C for 10 minutes. The electrophoresis was followed by washing the gels with distilled water for 30 minutes at room temperature. The fluorescent proteins in the unstained gels were visualized by using Dark Reader Blue transilluminator and by using UV imaging function of an AlphaImager gel documentation system (ProteinSimple). Gels were stained by PageBlue Protein Staining Solution (Thermo Fischer Scientific). Densitometry of gels was performed by using ImageJ 1.43 software (**Schneider et al., 2012**) and by GelAnalyzer 2010a program (www.gelanalyzer.com).

2.7 Calibration curve of the substrates

mTurquoise2 and mApple fluorescent protein-containing substrates were used for calibration. The purified substrates were dissolved in elution or cleavage buffer, and the solutions containing the substrates in different final concentrations were transferred into black half-area plates. Relative fluorescent intensities were measured by Biotek Synergy2 multimode plate reader using 590/35 nm excitation and 645/40 nm emission filters for mApple, while 400/10 nm excitation and 460/40 nm emission filters for mTurquoise2. Blank-corrected relative fluorescent intensity values (RFU) were plotted against the substrate concentration (mM). Linear regression was performed and the parameters of the fitted lines were determined by Microsoft Excel 2010 (Microsoft).

2.8 Kinetic and inhibition study using HIV-1 PR

Substrates were coated to Ni-NTA magnetic agarose beads as it was described in 2.4 section. The substrates coated to the beads were suspended in cleavage buffer, and then increasing amounts of the homogenous suspensions were measured into 2.0 ml Protein Lobind Micro-centrifuge tubes (Eppendorf). The tubes were applied to a Dynamag^{TM}-2 magnetic particle concentrator (Thermo Fischer Scientific, Invitrogen), the supernatant was removed and the beads were suspended in equal volume of cleavage buffer.

For determining the amount of the substrate attached to magnetic beads, substrate control samples were also prepared in the same way as the reaction samples except that the suspension was in elution buffer instead of cleavage buffer.

To determine kinetic time course for the release of proteolytic fragments, cleavage of His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and His_6-MBP-VSQNY↓PIVQ-mApple substrates (at 0.0016 mM and 0.0024 mM final concentrations, respectively) by HIV-1 PR (at 36.4 nM

final concentration) was followed by measuring fluorescence after 0, 20, 40 and 60 minutes incubation.

The purified HIV-1 PR was added to each reaction and incubated for 60 minutes at 37°C while continuously shaking at 600 rpm. Cleavage reactions were performed by using 36.4 nM HIV-1 PR (final, total enzyme concentration) and substrates containing HIV-1 matrix and capsid (MA/CA) cleavage site sequence (VSQNY↓PIVQ). Substrate blanks (reactions, where cleavage buffer was added instead of the enzyme) at each concentration points were also made to detect non-specific substrate dissociations. Reactions were terminated by separating the magnetic beads on Dynamag^{TM}-2 magnetic particle concentrator (Thermo Fischer Scientific, Invitrogen) and supernatants were measured as described in 2.7 section.

Initial velocity values (nMs^{-1}) were calculated from the blank-corrected RFUs using the slope of the substrate calibration curves in cleavage buffer and were plotted against the coated substrate concentration (mM) assessed from substrate control sample using the slope of the substrate calibration curve in elution buffer. Kinetic parameters were determined at less than 20% substrate turnover by Michaelis-Menten non-linear regression analysis using GraphPad Prism version 5.00 for Windows (GraphPad Software, La Jolla, California USA, www.graphpad.com).

His_6-MBP-VSQNY↓PIVQ-mTurquoise2 recombinant substrate (0.028 mM final concentration) was used to study inhibitory effect of amprenavir (ranging from 1nM to 1μM final total concentrations) on HIV-1 PR (at 36.4 nM final concentration). After the termination of the reactions the fluorescence of supernatants was measured. Initial velocity values (nMs^{-1}) were plotted against the logarithms of amprenavir concentrations (nM). Inhibitory effect at 50% initial velocity value (IC_{50}) was determined by fitting five parameter logistic curve on the data using GraphPad Prism version 5.00. Data of the present inhibitory measurement was applied for the determination of active enzyme concentration used for assessing kinetic parameters and inhibitory constant. The inhibitory constant (K_i) value was calculated from the IC_{50} value using the following equation: $K_i = (IC_{50}-[E]/2)/(1+[S]/K_m)$, where [E] and [S] are the concentrations of active enzyme (6.05 nM) and substrate, respectively, while K_m is the Michaelis constant.

2.9 Kinetic and pH dependence study of TEV PR

TEV PR stock solution, purified by the method of Kapust and its co-workers (**Kapust et al., 2001**), was a kind gift of dr. David S. Waugh (NCI-Frederick, USA).

His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and His_6-MBP-VSQNY↓PIVQ-mApple recombinant substrates were used to study the cleavage kinetics of TEV PR (the final concentration of the enzyme in the kinetic assays was 45.71 nM). Kinetic reactions were performed as described in 2.8 section, but incubation temperature was set to 30°C.

His_6-MBP-VSQNY↓PIVQ-mTurquoise2 recombinant substrate (at 0.03 mM final concentration) was used to study pH dependence of TEV PR. Substrates coated to Ni-NTA magnetic beads were purified as it was described in 2.4 section, but the pH of cleavage buffer was different (pH was set to 6.0, 6.5, 7.0, 7.5, 8.0, and 8.5). The final concentration of TEV PR was 91.42 nM in the reaction mixtures. Substrate blank samples (where enzyme buffer was added instead of the enzyme) were also prepared for each reaction to detect the dissociation of substrates from the beads at different pH values (the composition of the enzyme buffer was described previously in **Kapust et al., 2001)**. After the termination of reactions the fluorescence of supernatants was measured as it was described in 2.8 section. Initial velocity values ($nM^{-1}s^{-1}$) were plotted against pH.

2.10 Recycling of the magnetic beads

Ni-NTA magnetic beads (Qiagen) were used repeatedly during the experiments; the beads were regenerated before each use. Different buffers were used for all six washing steps of regeneration: (*i*) 0.5 M sodium-hydroxide and 0.05% Tween 20; (*ii*) 0.05% Tween 20; (*iii*) elution buffer; (*iv*) 0.05% Tween 20; (v) 100 mM $NiSO_4$ and 0.05% Tween 20; (vi) and 0.05% Tween 20. Each washing step was repeated five times. For long-term storage the beads were suspended into 30% ethanol containing 0.5% Tween 20 (pH 7.0).

2.11 HPLC-based protease assay

The protease assays were initiated by the mixing of 5 µl purified HIV-1 PR (dialyzed against 2X cleavage puffer, 430 nM final concentration), 10 µl 2X cleavage buffer, and 5 µl synthetic oligopeptide substrate (in 0.47-2.35 mM final concentrations) representing the naturally occurring MA/CA cleavage site of HIV1 PR (VSQNY↓PIVQ).

The reaction mixtures were incubated at 37°C for 10 minutes and stopped by the addition of 9 volumes of 1% trifluoroacetic acid (TFA). The samples were injected onto Nova-Pak C18 reversed-phase chromatography column (Waters Associates, Inc.) using an

automatic injector. Separation of substrates and cleavage products was performed by using acetonitrile gradient (0 to 100%) in water, in the presence of 0.05% TFA. We monitored the cleavage of peptides at 206 nm, followed by integration of the peak areas. Reactions were monitored at < 20% substrate hydrolysis.

Active enzyme concentration (22.05 nM) was determined the same way as kinetic measurements except that 2X cleavage buffer contained amprenavir as a potent inhibitor (final concentration ranging from 1 nM to 80 nM in reaction mixtures) at 0.47 mM final oligopeptide substrate concentration.

3. Results and discussion

3.1 Structure of the expression vectors

pDest-His_6-MBP-mApple and pDest-His_6-MBP-mTurquoise2 plasmids were prepared by using Gateway cloning technology based on the description of Tropea et al. (**2007**). We have designed N1 primer of the first PCR reaction in a way that besides amplifying the desired FP, it also codes for a so called 'cloning cassette', which is located between the TEV PR cleavage site and the FP. The schematic structure of a pDest-His_6-MBP-FP plasmid is illustrated in **Figure 1**.

Both pDest-His_6-MBP-mApple and pDest-His_6-MBP-mTurquoise2 plasmids carry the coding sequence of a His_6 affinity tag and a MBP fusion protein, followed by a TEV PR cleavage site (ENLYFQ↓G), the 'cloning cassette', and a C-terminal fluorescent protein (FP).

The nucleotide sequence coding for the desired proteolytic cleavage site can be inserted into the 'cloning cassette', which contains cleavage sites of PacI and NheI restriction endonucleases next to each other (see **Figure 1**). The plasmids linearized by PacI and NheI restriction enzymes have sticky ends and can be re-circularized after the addition of a linear dsDNA sequence flanked by PacI and NheI sticky ends. In our case, short and linear dsDNA sequences to be inserted were created by the attachment of complementary, *E. coli*-optimized (forward and reverse) oligonucleotide primers, which code for a cleavage site sequence of HIV-1 PR. The coding sequence of the HIV-1 PR cleavage site was inserted into pDest-His_6-MBP-mApple and pDest-His_6-MBP-mTurquoise2 plasmids by ligation.

It is important to note that the coding sequences of the fluorescent proteins are not in the open reading frame without the insertion of the properly designed linear dsDNA-s into the plasmids, therefore, the fluorescent proteins are translated only after a successful ligation.

3.2 The fluorescently labeled recombinant fusion protein substrates

Recombinant fusion proteins were expressed to contain an N-terminal hexahistidine affinity tag (His_6) and a maltose binding protein (MBP), followed by cleavage sites of TEV and HIV-1 proteases, respectively, while on the C-terminal end a fluorescent protein variant is fused to the cleavage site to be processed. The role of the His_6 affinity tag is to enable the recombinant proteins to be immobilized on metal-chelate surfaces, and also to make the purification of the proteins by immobilized metal affinity chromatography (IMAC) possible. The N-terminal MBP fusion protein enhances the water-solubility of the recombinant protein

substrate and improves its folding (**Kapust et al., 1999; Fox et al., 2003; Fox and Waugh, 2003**), while TEV protease cleavage site serves as an internal control cleavage site. Here we applied plasmids coding for C-terminal mApple or mTurquoise2 fluorescent proteins (**Shaner et al., 2008; Goedhart et al., 2012**). Compared to other similar protease assays (**Patel et al., 2001; Chaparro-Riggers et al., 2005**), our recombinant substrates contain an MBP fusion tag, as well, to provide a spacer preventing the possible influence of substrate immobilization on the enzyme kinetic parameters. Therefore, the cleavage sites of TEV and HIV-1 proteases are not located in the close proximity of the affinity tag responsible for the binding to the magnetic beads, and the flexibility of this MBP-FP interdomain linker provides sufficient accessibility of the cleavage sites in the immobilized substrates.

In our experiments, the recombinant protein substrates contained a cleavage site sequence in order to carry out activity measurements of HIV-1 PR. The inserted (9 residue-long) sequence represented the P5-P4' amino acid residues of the naturally occurring cleavage site between the matrix and capsid (MA/CA) of HIV-1 polyprotein (**Tözser et al., 1991**).

The recombinant substrates were successfully expressed in *E.coli* BL21(DE3) strain and the desired proteins were recovered from the cell lysate under native conditions for the purification and for the protease assay using Ni-NTA coated agarose beads.

The expressed fusion protein substrates in the cleared bacterial lysates were attached to Ni-NTA coated magnetic agarose beads. After the cleavage by HIV-1 PR or by TEV PR, the N-terminal part of the processed substrates remained attached to the magnetic beads and could be separated by a magnetic particle concentrator, while the fluorescent C-terminal part could be well-detected in the supernatant, based on which kinetic parameters can be calculated. The process of the assay is illustrated in **Figure 2**.

Ni-NTA beads were utilized not only for assay purposes, but also for purifying the substrates from the cleared bacterial lysates. Gel electrophoresis analysis of the purified substrates demonstrated that the fusion proteins had the expected molecular weight (~72 kDa) and purity for the assay. We have also verified that both TEV PR (data not shown) and HIV-1 PR cleaved the recombinant substrates efficiently, the cleavage products can be well-differentiated and visibly detected on the gel (**Figure 3**).

Effect of temperature on the stability of fusion protein substrates was also studied. We examined the effect of storage at 4°C and incubation at 37°C on substrate degradation. Stability assays were assessed by densitometric analyses of gels from repeated denaturing SDS-PAGE analyses (data not shown). We observed no significant substrate degradation

neither for storage at 4°C for at least two weeks nor for incubation at 37°C for 24 hours. Conformational stability of the His_6-MBP-VSQNY↓PIVQ-mApple recombinant substrate was assessed by tryptic digestion, as well (see methods and results in **Mótyán et al., submitted**). Results of this enzymatic analysis indicated proper folding of the substrate. We observed no decrease in band intensities for the Coomassie-stained protein products (data not shown) and for the UV image of mApple after in-gel renaturation (see section 3.3), which implied no degradation of the fusion partners while the accessible interdomain linker was cleaved by trypsin.

3.3 In-gel renaturation of fluorescent proteins

The purified recombinant protein substrates and cleavage products were analyzed by PAGE, using denaturing or non-denaturing conditions during the sample preparation, while polyacrylamide gels contained SDS in both cases. The samples, which were not denatured were treated with loading buffer containing no reducing agents and were not heat-treated before the electrophoresis. Loading buffer containing SDS and β-ME was added to the samples to be heat-treated for the denaturation of proteins.

The fluorescence of the bands in the case of non-denatured samples was readily detected by blue light transilluminator (**Figure 3A**) and by UV imaging (**Figure 3B**) using a gel documentation system after the electrophoresis. The presence of SDS in the gels makes the detection of the fluorescence of denatured fluorescent proteins impossible, furthermore, it interferes with Coomassie staining. Therefore, the electrophoresis was followed by washing the gels with distilled water to remove SDS and renature the proteins. We found that removal of the SDS from the gels made the detection of fluorescence possible by using blue light transilluminator and UV imaging (**Figure 3A and 3B**).

To our knowledge, protocol for the detection of fluorescent proteins after a denaturing SDS-PAGE analysis (including heat-treatment) by washing out the SDS from the gel has not been published till now.

The fluorescence of denatured recombinant protein substrates and products were successfully detected in the gel by UV imaging after the removal of SDS, because the washing step was found sufficient for at least partial renaturation of the proteins. The advantage of the application of protein denaturation/renaturation in this protease assay is that the cleavage products can be separated from each other based on their molecular weight, and their native charge or shape does not affect their migration during electrophoresis. Besides the

differentiation based on protein size, the appropriate bands can be identified based on their fluorescence, as well. The protein size-based identification could be relevant if the molecular weights of other cleavage products (or contaminants) having no fluorescence closely resemble to those of the fluorescent products. This is represented in **Figure 3C**, where the cleavage products are not separated fully from each other in the case of non-denaturing electrophoresis, while can be differentiated in renatured samples after Coomassie staining of the gel (**Figure 3C**). In the case of multiple cleavages within the inserted target sequence, the cleavage products having different molecular weights may be also identified by denaturing gel electrophoresis, while cannot be differentiated purely based on measuring the fluorescence of the supernatant.

We have investigated the resolution ability of the polyacrylamide gel electrophoresis to separate cleavage products having similar molecular weights. His_6-MBP-VSQNY↓PIVQ-mApple recombinant protein substrate was digested by HIV-1 and TEV proteases, and the proteolytic fragments were separated by SDS-PAGE using a 14% polyacrylamide gel. The ~1.5 kDa differences between the cleavage products, which showed only 15 amino acid residue difference in their lengths, were clearly visible in the Coomassie-stained gel (**Figure 3D**). This implies that SDS-PAGE may be indicative of alternative cleavages within the substrates, and the proteolytic fragments showing only slight differences in their molecular weights can be differentiated. As the resolution is strongly dependent on the type of the applied gel, theoretically it can be further improved by increasing the gel concentration or by applying tricine-SDS-PAGE.

It is important to note that differences were observed in the renaturation of mApple and mTurquoise2 fluorescent proteins. mTurquoise2 showed substantially better renaturation ability compared to mApple, when detected by blue light transilluminator and UV imaging (**Figure 3A**). While band intensity of renatured mTurquoise2 compared to the non-denatured was at least 50% based on densitometry, mApple showed only ~10% fluorescence intensity but was detectable in the gel under UV light (**Figure 3B**). It has already been reported by Gross et al. **(2000)** that the acylimine bond of the dsRed chromophore is irreversibly degraded upon denaturation. The acylimine bond, being responsible in part for the red shift of mApple (**Shu et al., 2006**) may be also disrupted upon denaturation and this change can be possibly responsible for the lower renaturation ability of mApple. While detailed analysis of the changes induced by denaturation may help understanding the differences between the

sensitivity of the fluorescent proteins, the investigation of renaturation abilities was out of the scope and extent of this study.

3.4 Substrate calibration curves

We have tested whether the fluorescence intensity of the purified substrates is proportional to their concentration. Relative fluorescence intensities of substrates in either elution or cleavage buffers were measured to prepare the calibration curves. Based on the calibration we found that the relative fluorescence intensities are directly proportional to the substrate concentration in the range applied in our experiments (**Figure 4**).

3.5 Kinetic and inhibition study of HIV-1 PR

Although published *in vitro* protease assay systems (**Chaparro-Riggers et al., 2005**; **Patel et al., 2001; Askin et al., 2011; Zhou et al., 2014**) may be adapted to several proteases, to our knowledge, the applications of fluorescent protein-based assays for the measurement of HIV-1 PR activity have not been published so far.

In order to demonstrate the suitability of the described protease assay system, we performed enzyme kinetic measurements by using substrates, which contain the HIV-1 MA/CA cleavage site sequence. Kinetic parameters summarized in **Table 1A** show that k_{cat}/K_m values for HIV-1 PR measured on mApple and mTurquoise2 substrates are practically identical, while the individual k_{cat} and K_m values differed substantially. Representative graphs for non-linear regression analyses of measured data are shown in **Figure 5A and B,** and **Figure 5C and D** graphs show that kinetic time courses are linear for the fluorescent cleavage products. For comparative purposes we have also determined the kinetic parameters for HIV-1 PR using an oligopeptide substrate **(Table 1B**), where folding efficiency of HIV-1 PR was calculated to be 5.13 %. The k_{cat}/K_m values determined for HIV-1 PR by oligopeptide substrate closely resemble that of measured by the recombinant protein substrates, while the individual kinetic parameters were substantially higher in the peptide assay. These differences might be accounted for the effect of different substrate types, like extended binding site and effect of neighbouring protein domains on the cleavage site flexibility. Our previously published much higher specificity constant (45.3 $mM^{-1}\,s^{-1}$) for the oligopeptide substrate representing the same cleavage site was determined under very different assay conditions, including the different buffer systems, which have different pH (pH 5.6) and ionic strength (2

M) (**Tözsér et al., 1991**) compared to the recombinant protein substrate-based assay (pH 7.0 and 300 mM NaCl concentration).

It is important to note, that the recombinant fusion protein-based assays were performed at pH 7.0, because substrates attachment to the bead-surface is pH dependent and most effective over pH 6.0. However, the applied pH is higher than in the case of synthetic oligopeptide substrate-based HPLC assays (pH 5.6) (e.g. **Fehér et al., 2006; Eizert et al., 2008**), the HIV-1 PR activity was successfully measured at neutral pH on the recombinant fusion protein substrates. It should be noted that in our assay relatively low salt concentration (300 mM NaCl) was applied compared to the high salt concentration (2 M or more), which was previously found to be optimal for retroviral proteases and was used in common HPLC-based assays (e.g. **Bagossi et al., 2004; Fehér et al., 2006**).

The HIV-1 protease inhibitor amprenavir was used at increasing concentrations to test its inhibitory effect on the cleavage of His_6-MBP-VSQNY↓PIVQ-mTurquoise2 fusion protein substrate by HIV-1 PR. The sigmoid dose-response curve (represented in **Figure 6**) was used to determine IC_{50} (21.43 nM) and K_i values (6.75 nM).

Our results demonstrate that this substrate system is suitable to study the inhibitory potential of protease inhibitors on the enzyme activity.

3.6 Kinetic and pH dependence study of TEV PR

In order to examine the internal control TEV PR cleavage site of the fusion protein substrates, kinetic measurements by TEV PR were also performed on His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and His_6-MBP-VSQNY↓PIVQ-mApple substrates.

Values summarized in **Table 1A** show that kinetic parameters determined on mApple and mTurquoise2 substrates are very similar, and the k_{cat}/K_m value is also similar to what was determined for TEV PR using an oligopeptide substrate TENLYFQ↓GGTRR (3.08 ± 0.67 mM^{-1} s^{-1}; **Kapust et al., 2002**), while both the K_m and k_{cat} values are substantially lower than those found in the peptide-based assay (K_m = 0.087 ± 0.017 mM; k_{cat} = 0.27 ± 0.03 s^{-1}; **Kapust et al., 2002**). Differences may be attributed to the different substrate types as detailed for HIV PR cleavage, furthermore, the buffer systems and ionic strengths were also different in the oligopeptide- and the recombinant protein-based protease assays (400 mM and 300 mM NaCl, respectively), furthermore, no reducing agents were added to the buffer in the herein described protease assay, while activity measurements of TEV PR have been performed in the presence of DTT (**Kapust et al., 2002**).

TEV PR was found previously to have no activity at or below pH 5.0 but is active over a pH range between 6.0 and 9.0 (**Parks et al., 1995**), therefore, this enzyme was chosen to demonstrate that the method is suitable for the determination of the pH dependence of enzymes that prefer alkaline reaction conditions (**Figure 7**). TEV PR activity was measured in the 6.0-8.5 pH range, and we found no statistically significant differences between the activities measured at pH 6.5, 7.5, 8.0, and 8.5 compared to the highest activity observed at pH 7.0. These results are in agreement with the literature data (**Parks et al., 1995**) and imply an optimal pH range for TEV PR.

Standard deviation (SD) at pH 6.0 was found to be relatively high compared to SDs of the results at other pHs (data not shown). This phenomenon may be caused by the nature of the interaction between the Ni-NTA agarose bead and the His_6 fusion tag of a protein substrate, which significantly weakens at or under pH 6.0 (in agreement with the manufacturer's specifications). Our results also showed substantial substrate dissociation at pH 6.0, while only weak dissociation was observed at higher pH (**Figure 8**). Spontaneous substrate dissociation was considered during the evaluation of kinetic measurements and values measured for the enzyme reactions have been corrected by the fluorescence of substrate blank samples.

3.7 Advantages and limitations of the assay system

Due to the flexibility of the 'cloning cassette', the expression system allows easy generation of a substrate library, the desired oligonucleotide sequences can be inserted into the 'cloning cassette' simply by a one-step PCR reaction. The expression construct was designed to make the ligation efficiency easy-to-follow by a simple transilluminator, because the FP is only translated if the insertion of the cleavage site is successful.

Besides the mApple and mTurquoise2, other fluorescent proteins (e.g. mCherry, green fluorescent protein, etc.) can be also attached to the fusion protein by the modification of the expression construct. Since 3' and 5' ends of the DNA sequences of the most commonly used fluorescent proteins are highly similar, numerous plasmid variants can be created by using only one pair of N1 and C oligonucleotide primer. Fluorescent proteins ensure highly sensitive fluorescent detection in the designed assay and - as an extra advantage - it makes the recombinant fusion substrate easy-to-follow during expression, purification and working procedures. Fluorescent proteins that most suit the experimental purposes and the given instrumentation can be chosen (**Davidson et al., 2009**).

Although protease assay systems based on the use of recombinant protein-based substrates (consisting of an affinity tag, a proteolytic cleavage site of interest, and a fluorescent protein) have already been developed (**Patel et al., 2001; Chaparro-Riggers et al., 2005**; **Askin et al., 2011**), our system may provide a novel tool for protease research by combining and also improving the advantages of the already developed methods by offering a true separation-based assay that enables the determination of protease kinetic parameters (k_{cat}, K_m) supported by a detailed substrate quantification procedure (linearity of the fluorescence versus substrate concentration is examined in a wide concentration range, also considering the difference in the relative fluorescence in elution and cleavage buffer) in a low volume, HTS-compatible format, utilizing a fusion protein substrates equipped with MBP to enhance protein solubility, built-in TEV PR cleavage site to serve as a control cleavage site, and with monomeric FPs to avoid substrate aggregation. Moreover, for a simple and fast demonstration of cleavage of the desired substrate with a specific enzyme, the fluorescent substrates and the generated products can be visualized by gel electrophoresis (**Figure 3**). Our observations were in agreement with the previous observations of Saeed and its co-workers, who found that the fluorescent proteins exposed to denaturants show differences in their stability despite the close relationships (**Saeed et al., 2009**). Therefore, abilities of proteins for renaturation need to be considered in the experimental design, fluorescent proteins with higher stability need to be chosen or in-gel renaturation needs to be further optimized to make the detection of less-stable proteins sensitive enough.

The herein described recombinant fusion protein substrate-based protease assay has the advantage that it can be effortlessly adapted to HTS environment (**Inglese et al., 2007**) due to the followings: (*i*) the reaction can be well performed using small volume of reagents (in our case total reaction volume was under 70 µl); (*ii*) no built-in factors of the system limit the lengths of the assay procedures, it depends only on the incubation time; (*iii*) the assay procedure itself consists only of 6-7 short steps that are fully automation compatible; (*iv*) no special instrumentation is needed, which makes the assay especially cost-efficient.

Due to the nature of the fluorescence read out, the assay is highly sensitive, and despite of the fact that it is an endpoint method, the system also offers good relative estimation of the catalytic efficiency. Other attractive feature of the assay is that simple and safe operation is ensured during the whole process.

Since both the substrates and the products have similar fluorescence due to the presence of the fluorescent protein tag, they cannot be differentiated in solution by

fluorimetric measurements, therefore, in our assay system the substrates are attached to magnetic beads to make the separation of the cleavage fragments (present in the supernatant) from the uncleaved proteins possible, likewise in the case of other similar assays (**Patel et al., 2001; Chaparro-Riggers et al., 2005**). On the other hand it is also possible to study proteolytic cleavage of the previously purified recombinant substrates by in-solution digestion, and in this case uncleaved substrates and the cleavage fragments can be differentiated by electrophoretic separation and molecular weight-based identification.

We have also tested the feasibility of the solution phase-extraction method for the removal of the substrates and processed N-terminal fragments from the reaction mixtures. Purified His_6-MBP-VSQNY↓PIVQ-mApple substrate was incubated with the Ni-NTA coated magnetic beads, and the dependence of substrate binding on bead concentration and incubation time was studied (see methods and results in **Mótyán et al., submitted**). It was found that the substrates are able to be withdrawn from the solution by the addition of Ni-NTA magnetic beads, and increasing bead concentration or incubation time also resulted in elevated attachment of the substrates to the beads. This implies that the magnetic beads may be useful for the removal of uncleaved substrates from the reaction mixtures in the case of an in-solution digestion assay, however, the optimization of the binding conditions (bead concentration and incubation time) is necessary. Avoiding the use of long incubation time is advisable, and high bead concentrations could possibly make handling the mixtures difficult especially in a low volume format. It is also important to note that in order to set an in-solution protease assay, beyond the optimization of substrate removal, a proper method that do not interfere with the magnetic beads and with the fluorescence of the recombinant proteins, is also needed to be optimized to stop the cleavage reaction.

Two methods widely applied for stopping enzymatic reactions have been tested by examining the bead-binding and fluorescence properties of His_6-MBP-VSQNY↓PIVQ-mApple substrate after trichloroacetic acid (TCA)- and heat-treatments (see methods and results in **Mótyán et al., submitted**). As it was expected, TCA-treatment (using TCA in 5% final concentration) abolished fluorescent property of the substrate and prevented effective binding of the substrate to the magnetic beads. Interestingly, heat-treatment (incubation at 95°C for 10 minutes) has not caused the total loss of fluorescent property of the substrate, approximately 50% decrease in the signal intensity was observed, and similarly to TCA-treatment, it also substantially inhibited bead attachment. The above mentioned methods were

not found to be useful for stopping the reaction due to their adverse effects on the fluorescence properties of the substrates.

Some more features that can make the system more attractive is that based on our observations one calibration curve suits for a set of substrates (data not shown), the measurements are characterized by low level of impurities and due to the strong nature of His_6-tag and Ni-NTA interactions the background of the measurement due to non-specific substrate leakage is especially low.

However, just like any other analytical and screening methods this protease assay system also has some of its limitations. For example, the reaction and buffer conditions must be adjusted to the optimal conditions of Ni-NTA magnetic beads, and the effective attachment of the substrates to the bead-surface is pH dependent. According to the manufacturer's handbook, monomer proteins start to be released from the bead surface at pH 5.9, whereas multimers are released at pH 4.5. Although the presented assay works reliably over pH 6.0, in some cases lower pHs may also be applied, however, the rate of spontaneous substrate dissociation at each reaction must be followed by appropriate control reactions and must be considered at the evaluation of the results. We also observed increased spontaneous dissociation of substrates at pH 6.0, while increasing the pH led to weaker dissociation (**Figure 8**).

In our experiments purified enzymes were applied in cleavage reactions, but in some cases the use of unpurified protease solutions may also be desired. Working with crude cell lysates needs to be carefully optimized, the nature of the protease of interest and the sample origin (e.g. bacterial, eukaryotic) is also required to be considered. To avoid unwanted cleavages by other proteases which may cause distortion of the results, the application of protease inhibitor cocktails may also be necessary. An in silico prediction (performed by the PeptideCutter module of ExPASy, accessible at http://web.expasy.org/peptide_cutter) showed the presence of possible cleavage sites for several proteases within the recombinant protein substrate sequences. This implies that addition of a crude cell lysate to the purified recombinant substrate may cause cleavages at multiple sites. However, it is important to note that we have tested the susceptibility of a recombinant substrate towards digestion by trypsin, and our results showed proper folding of the expressed substrates, thus no degradation of the fusion partners was observed. This implies that the MBP and fluorescent proteins may be not cleaved efficiently by endogenous proteases, but the interdomain linkers are accessible and

can be sensitive towards proteolysis by endogenous proteases. In the case of total cell lysate samples the in-gel detection may be problematic in Coomassie-stained gels because the bands of substrate and cleavage products are buried with the endogenous protein bands, but according to our denaturing SDS-PAGE analyses the fluorescent proteins can be properly detected in gels by UV imaging (data not shown). While the recombinant substrates may be useful only for the detection of protease activities of unpurified proteases, a great disadvantage of this application is that it is not useful for determination of enzyme kinetic parameters.

Our studies showed stability of the recombinant substrates, which implies that longer incubation times may be also used in the designed protease assay system due to the stability of the recombinant substrates at 37°C for at least 24 hours. Conformational stability of the substrates was proved by tryptic digestion, proper folding of the recombinant protein prevented fragmentation the fusion partners.

Future users may also need to pay special attention to the type of the applied fluorescent protein tags in the reaction, since fluorescent proteins have the tendency to produce dimers and some of them may be unstable, therefore, monomer and stable variants are preferred (**Shaner et al., 2005; Shaner, 2014**).

4. Conclusions

Here we describe the development and application of a recombinant fusion protein substrate-based protease assay.

Plasmids coding for the backbone of the recombinant, fluorescent protein-coding substrates were created using Gateway cloning technology. Each expression construct comprises the coding sequences of a His_6 and an MBP N-terminal fusion tag, a TEV PR processing site, a 'cloning cassette', and a C-terminal fluorescent protein. Linear DNA sequences coding for the cleavage sequences to be examined can be inserted into the designed 'cloning cassettes' of the plasmids by ligation. The whole fusion protein substrates can be expressed in *E. coli* BL21(DE3) cells and purified by Ni-NTA-coated magnetic agarose beads. The bead-attached recombinant protein substrates containing the sequences to be cleaved are processed by the protease of interest. The reaction can be terminated by the removal of the N-terminal products by using a magnetic particle separator, and based on the detected fluorescence of the released C-terminal fluorescent products the kinetic parameters can be determined. Ni-NTA coated magnetic agarose beads facilitate the easy purification of the intact protein substrates for other analytical purposes, e.g.: substrate calibration, protein content determination or PAGE. Due to the fluorescent properties of the protein substrates the whole working procedure is especially easy to follow visually.

The protease assay was successfully applied for the activity measurements of a HIV-1 PR and TEV PR. The assay system was found to be useful to determine enzyme kinetic parameters, to determine the inhibitory potential of amprenavir on HIV-1 PR, and the pH dependence of TEV PR, as well.

Our protease assay system shares some similarities with previously developed methods (**Patel et al., 2001; Chaparro-Riggers et al., 2005; Askin et al., 2011**), that are also based on the use of recombinant protein substrates containing an affinity tag (for substrate immobilization) and a fluorescent fusion tag (for detection) besides the proteolytic cleavage sites. The bead-attached substrates and the protease cleavage-released products can be quantified by fluorimetric measurements. However, these recombinant protein substrate-based assays are theoretically suitable for kinetic measurements, only one of these assays (**Chaparro-Riggers et al., 2005**) have been used for the determination of kinetic parameters for the studied enzyme. Similarly to the substrate designed by Patel et al. **(2001)**, our substrates also contain a control cleavage site for TEV protease.

Our assay format integrates most of the advantages of previously mentioned fusion protein-based platforms, furthermore, important assay elements of these methods are applied in an improved format. Furthermore, despite the above mentioned similarities, our assay was designed to contain improvements compared to the previously described systems. An advantage of our system is the possibility for the determination of the kinetic parameters supported by a detailed substrate quantification procedure (linearity of the fluorescence versus substrate concentration is examined in a wide concentration range, also considering the difference in the relative fluorescence in elution and cleavage buffer), which helps the generation of more reliable results. Moreover, a completion of the recombinant substrates has been also performed by the insertion of an MBP fusion protein prior to the protease cleavage sites to improve the folding efficiency and solubility of the substrates, further expanding the potential use of our assay system. Introduction of MBP was expected to minimize possible effects of immobilization on substrate cleavage due to the increased distance between the cleavage sites and the affinity tag. However, adverse effects of immobilization (due to the proximity of the cleavage site sequences to the affinity tags) have not been described for the similar assays (**Patel et al., 2001; Chaparro-Riggers et al., 2005**). Another advantage of this system is the lowered volume format (final volume of reaction mixtures is ≤70 µl), which altogether with a simplified assay procedure makes the assay compatible with HTS systems. By the application of the recently engineered monomeric mApple and mTurquoise2 fluorescent proteins the potential aggregation of substrates can be minimized.

While most HIV protease assays are based on the use of oligopeptide substrates, for studying the external binding sites of the HIV-1 PR and other retroviral or retroviral-like proteases elongated substrates are necessary. The application of the recombinant fusion protein substrates provide good alternative for the synthetic oligopeptide substrates. One of the main advantages of the developed protease assay is that both the full-length substrates and the cleavage products can be readily detected visually or by blue transilluminator or UV imaging system after PAGE analysis. We found that the in-gel UV detection of the recombinant protein substrates and cleavage products containing a fluorescent protein tag is also possible after the in-gel renaturation of the proteins after a denaturing SDS-PAGE (including heat-treatment of the samples). Stability of the substrates provides their storage at 4°C and makes the application of longer incubation times at 37°C possible.

Due to the flexibility of the assay system, wide variety of the sequences can be easily ligated into the ‘cloning cassettes’ of the expression plasmids. The cost-efficient and

expeditious generation of the fluorescent fusion substrates harbouring the cleavage site of interest, makes the system especially attractive for the generation of substrate libraries, by which the method can be a good alternative tool of protease specificity measurements (e.g. similar to ones described by **Eizert et al. 2008**), mutagenesis studies, or due to the HTS-compatible format of the entire assay procedure (from bacterial cell disruption to the determination of the kinetic parameters), it could be also effectively utilized for industrial protease inhibitor screening and/or antiviral drug development. Moreover, in principle, by mixing substrates bearing different C-terminal fluorescent tags, the assay may be suitable for the examination of competitive proteolysis of a protease of interest on different cleavage sites, in one reaction. Further investigation and optimization of the assay for this future application is also in the scope of our laboratory.

5. Acknowledgement

This work was supported in part by the GINOP-2.3.2-15-2016-00044 "PHARMPROT teaming" project. We would like to say special thanks for Gedeon Richter Plc., especially to Dr. Zoltán Urbányi for allowing Beáta Bozóki's work in the Department of Biochemistry and Molecular Biology as guest researcher. We must also acknowledge Dr. Attila Bácsi (UD, Department of Immunology) for his help in the fluorescent measurements. We are grateful for all the members of Laboratory of Retroviral Biochemistry, especially to Szilvia Janics-Pető for her kind and immediate help in routine laboratory work. The authors are grateful to Dr. István Szatmári from Department of Biochemistry and Molecular Biology for critically reading the manuscript.

Author Contribution Statement

Beáta Bozóki: planned and performed experiments, analyzed data, wrote the paper.

Ferenc Tóth: performed experiments, analyzed data, wrote the paper.

Lívia Gazda: performed experiments, analyzed data, wrote the paper.

Márió Miczi: performed experiments, analyzed data, wrote the paper.

János András Mótyán: planned and performed experiments, analyzed data, wrote the paper.

József Tőzsér: conceived and managed the study, planned experiments, analyzed data, wrote the paper.

All authors read and approved the final version of the submitted manuscript.

The authors declare that there is no conflict of interest.

Authors declare that the work described in this manuscript has not been published previously and it is not under consideration for publication elsewhere.

Color should be used for Figure 2. and 3. in print.

6. References

Askin SP, Morin I, Schaeffer PM. **Development of a protease activity assay using heat-sensitive Tus-GFP fusion protein substrates**. Anal Biochem. 2011; 415(2): 126-33.

Aoki T, Tsuchida S, Yahara T, Hamaue N, **Novel assays for proteases using green fluorescent protein-tagged substrate immobilized on a membrane disk**. Anal. Biochem. 2008; 378: 132–137.

Bagossi P, Kádas J, Miklóssy G, Boross P, Weber IT, Tözsér J. **Development of a microtiter plate fluorescent assay for inhibition studies on the HTLV-1 and HIV-1 proteinases.** J Virol Methods. 2004; 119(2):87-93.

Boross P, Bagossi P, Copeland TD, Oroszlan S, Louis JM, Tözsér J**. Effect of substrate residues on the P2' preference of retroviral proteinases**. Eur J Biochem. 1999; 264(3):921-9.

Branchini BR, Rosenberg JC, Ablamsky DM, Taylor KP, Southworth TL, Linder SJ. **Sequential bioluminescence resonance energy transfer-fluorescence resonance energy transfer-based ratiometric protease assays with fusion proteins of firefly luciferase and red fluorescent protein**. Anal Biochem. 2011; 414(2):239-45.

Chaparro-Riggers JF, Breves R, Michels A, Maurer KH, Bornscheuer U. **A GFP-based assay for the determination of hydrolytic activity and substrate specificity of subtilisins under washing conditions**. J. Mol. Catal. B: Enzym. 2005; 35: 74-77.

Choe Y, Leonetti F, Greenbaum DC, Lecaille F, Bogyo M, Brömme D, Ellman JA, Craik CS. **Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities.** J Biol Chem. 2006; 281(18):12824-32.

Davidson MW, Campbell RE. **Engineered fluorescent proteins: innovations and applications.** Nat Methods. 2009; 6(10):713-17.

Eizert H, Bander P, Bagossi P, Sperka T, Miklóssy G, Boross P, Weber IT, Tözsér J. **Amino acid preferences of retroviral proteases for amino-terminal positions in a type 1 cleavage site**. J Virol. 2008; 82(20):10111-7.

Fehér A, Boross P, Sperka T, Miklóssy G, Kádas J, Bagossi P, Oroszlan S, Weber IT, Tözsér J. **Characterization of the murine leukemia virus protease and its comparison with the human immunodeficiency virus type 1 protease**. J Gen Virol. 2006; 87(Pt 5):1321-30.

Fox JD, Routzahn KM, Bucher MH, and Waugh DS. **Maltodextrin binding proteins from diverse bacteria and archaea are potent solubility enhancers**. FEBS Letts. 2003, 537, 53-57.

Fox JD and WaughDS. **Maltose-binding protein as a solubility enhancer**. Methods Mol. Biol. 2003, 205, 99-117.

Goedhart J, von Stetten D, Noirclerc-Savoye M, Lelimousin M, Joosen L, Hink MA, van Weeren L, Gadella TW Jr, Royant A. **Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%**. Nat Commun. 2012; 3:751.

Gross LA, Baird GS, Hoffman RC, Baldridge KK, Tsien RY. **The structure of the chromophore within DsRed, a red fluorescent protein from coral**. Proc Natl Acad Sci U SA. 2000; 97(22):11990-5.

Hassiepen U, Eidhoff U, Meder G, Bulber JF, Hein A, Bodendorf U, Lorthiois E, Martoglio B. **A sensitive fluorescence intensity assay for deubiquitinating proteases using ubiquitin-rhodamine110-glycine as substrate**. Anal Biochem. 2007; 371(2):201-7.

Inglese J, Johnson RL, Simeonov A, Xia M, Zheng W, Austin CP, Auld SD. **High-throughput screening assays for the identification of chemical probes.** Nat Chem Biol. 2007; 3(8):466-79.

Jadhav A, Ferreira RS, Klumpp C, Mott BT, Austin CP, Inglese J,Thomas CJ, Maloney DJ, Shoichet BK, Simeonov A. **Quantitative Analyses of Aggregation, Autofluorescence, and Reactivity Artifacts in a Screen for Inhibitors of a Thiol Protease.** J Med Chem. 2010; 53(1): 37-51.

Kapust RB and Waugh DS. **Escherchia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused**. Protein Sci. 1999: 8, 1668-1674.

Kapust RB, Tözsér J, Fox JD, Anderson DE, Cherry S, Copeland TD, Waugh DS. **Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency.** Protein Eng. 2001; 14(12):993-1000.

Kapust RB, Tözsér J, Copeland TD, Waugh DS. **The P1' specificity of tobacco etch virus protease.** Biochem Biophys Res Commun. 2002; 294(5):949-955.

Kostallas G, Löfdahl PÅ, Samuelson P. **Substrate profiling of tobacco etch virus protease using a novel fluorescence-assisted whole-cell assay**. PLoS One. 2011; 6 (1):e16136.

Levine LM, Michener ML, Toth MV, Holwerda BC. **Measurement of specific protease activity utilizing fluorescence polarization**. Anal Biochem. 1997; 247(1):83-8.

Louis JM, Clore GM, Gronenborn AM. **Autoprocessing of HIV-1 protease is tightly coupled to protein folding.** Nature Structural Biology, 1999; 6(9):868-875.

Mahalingam B, Louis JM, Reed CC, Adomat JM, Krouse J, Wang YF, Harrison RW, Weber IT. **Structural and kinetic analysis of drug resistant mutants of HIV-1 protease**. Eur J Biochem. 1999; 263(1):238-45.

Mahalingam B, Louis JM, Hung J, Harrison RW, Weber IT. **Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes**. Proteins. 2001; 43(4):455-64.

Matayoshi ED, Wang GT, Krafft GA, Erickson J. **Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer**. Science. 1990; 247(4945):954-8.

Mótyán JA, Miczi M, Bozóki B, Tözsér J. **Data supporting Ni-NTA magnetic bead-based fluorescent protease assay using recombinant fusion protein substrates**. Data in Brief. Submitted.

Parks TD, Howard ED, Wolpert TJ, Arp DJ, Dougherty WG. **Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form.** Virology. 1995; 210(1): 194-201.

Patel D, Frelinger J, Goudsmit J, Kim B. **In vitro assay for site-specific proteases using bead-attached GFP substrate.** Biotechniques. 2001; 31(5): 1194, 1196, 1198 passim.

Saeed IA, Ashraf SS. **Denaturation studies reveal significant differences between GFP and blue fluorescent protein.** Int J Biol Macromol. 2009; 45(3):236-41.

Schneider CA, Rasband WS, Eliceiri KW. **NIH Image to ImageJ: 25 years of image analysis.** Nat Methods. 2012; 9(7):671-5.

Shaner NC, Steinbach PA, Tsien RY. **A guide to choosing fluorescent proteins**. Nat Methods. 2005; 2(12):905-9.

Shaner NC, Lin MZ, McKeown MR, Steinbach PA, Hazelwood KL, Davidson MW, Tsien RY. **Improving the photostability of bright monomeric orange and red fluorescent proteins.** Nat Methods. 2008; 5(6):545-51.

Shaner NC. **Fluorescent proteins for quantitative microscopy: important properties and practical evaluation**. Methods Cell Biol. 2014; 123:95-111.

Shu X, Shaner NC, Yarbrough CA, Tsien RY, Remington SJ. **Novel chromophores and buried charges control color in mFruits.** Biochemistry. 2006; 45(32):9639-47.

Tözsér J, Bláha I, Copeland TD, Wondrak EM, Oroszlan S. **Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in Gag and Gag-Pol polyproteins.** FEBS Lett. 1991; 281(1-2):77-80.

Tropea JE, Cherry S, Nallamsetty S, Bignon C, Waugh DS. **A generic method for the production of recombinant proteins in Escherichia coli using a dual hexahistidine-maltose-binding protein affinity tag**. Methods Mol Biol. 2007; 363:1-19.

Zhang G. **Protease Assays.** In: Sittampalam GS, Coussens NP, Brimacombe K, et al., (Eds.), Assay Guidance Manual [Internet], Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2012.

Zhou C, Yan Y, Fang J, Cheng B, Fan J. **A new fusion protein platform for quantitatively measuring activity of multiple proteases**. Microb Cell Fact. 2014; 13(1):44.

Figure legends

Figure 1. Schematic representation of the pDest-His$_6$-MBP-FP expression vector coding the recombinant fusion protein substrates. Nucleotide sequence of the 'cloning cassette' is also indicated, recognition sequences of PacI and NheI endonucleases are underlined, while restriction endonuclease cleavage sites are indicated by asterisks. His$_6$: hexahistidine tag; MBP: maltose binding protein; TEV: tobacco etch virus; FP: fluorescent protein. Arrow shows cleavage position in TEV protease cleavage site (ENLYFQ↓G).

Figure 2. Principle of the fluorescent protease assay. The cleavage of the recombinant substrate by HIV-1 PR is illustrated on the example of His$_6$-MBP-VSQNY↓PIVQ-mTurquoise2 substrate. C-terminal part of the substrate (containing the fluorescent tag) is released by the proteolysis into the supernatant and can be detected by fluorimetry.

Figure 3. In-gel renaturation of recombinant fluorescent protein substrate and cleavage products. Purified His$_6$-MBP-VSQNY↓PIVQ-FP proteins were used as substrates for cleavage by HIV-1 PR. Both non-denatured and denatured/renatured samples were analyzed by SDS-PAGE. Sample denaturation was performed by heat-treatment of the samples at 95°C in the presence of SDS and β-ME. **A)** The unstained gels were illuminated by blue light transilluminator before (left panel) and after (right panel) washing the gel, as well. Fluorescent bands of the molecular weight standard are also indicated. **B)** Bands of the recombinant substrates are also shown and represent more efficient renaturation of mTurquoise2 fluorescent protein compared to mApple, as it was visualized under UV light after washing the SDS out from the gel. **C)** In the case of denaturation, the two cleavage products of His$_6$-MBP-VSQNY↓PIVQ-mTurquoise2 substrate were separated from each other based on their molecular weight, and can be differentiated in the Coomassie-stained gel, while less prominent separation was observed for the non-denatured sample. **D)** His$_6$-MBP-VSQNY↓PIVQ-mApple substrate was cleaved by HIV-1 or TEV proteases and the cleavage fragments were separated on 14% polyacrylamide gel after denaturing the samples, the gel was stained by Coomassie dye. The proteolytic cleavage sites are underlined within the

sequences of MBP-mApple interdomain linkers (arrows indicate cleavage positions), and calculated molecular weights of the substrate and the products are also indicated.

Figure 4. Substrate calibration curves of relative fluorescence intensity units versus substrate concentration. Linear regression analysis was used to evaluate results of the calibration and determine R^2 and slope values. Graphs show the calibration curves of His_6-MBP-VSQNY↓PIVQ-mTurquoise2 (**A, B**) and His_6-MBP-VSQNY↓PIVQ-mApple (**C, D**) substrates in cleavage (**A, C**) and elution buffers (**B, D**).

Figure 5. Michaelis–Menten plots of velocity versus substrate concentration. Graphs show the results of cleavage of His_6-MBP-VSQNY↓PIVQ-mApple (**A**) and His_6-MBP-VSQNY↓PIVQ-mTurquoise2 (**B**) substrates by HIV-1 PR. PIVQ-mApple (**A**) and PIVQ-mTurquoise2 (**B**) fluorescent cleavage products were measured by a fluorescence plate reader in the supernatants of the cleavage reactions. Non-linear regression analysis was used to evaluate the results of the protease assays and determine v_{max} and K_m values. Representative time courses are also shown for His_6-MBP-VSQNY↓PIVQ-mApple (**C**) and His_6-MBP-VSQNY↓PIVQ-mTurquoise2 (**D**) substrates in the case of cleavage by HIV-1 PR. Concentrations of PIVQ-mApple (**C**) and PIVQ-mTurquoise2 (**D**) fluorescent cleavage products have been calculated and plotted against time.

Figure 6. Inhibition of HIV-1 PR by amprenavir. His_6-MBP-VSQNY↓PIVQ-mTurquoise2 fusion protein was used as substrate for inhibition measurement.

Figure 7. Dependence of TEV PR activity on pH. Activity of TEV PR was measured at different pH by using His_6-MBP-VSQNY↓PIVQ-mTurquoise2 fusion protein as substrate.

Figure 8. Spontaneous dissociation of substrates from the magnetic beads is dependent on pH. Spontaneous substrate dissociation from the Ni-NTA agarose bead surface was examined in TEV PR pH dependence study by using His_6-MBP-VSQNY↓PIVQ-mTurquoise2 fusion protein as substrate.

Table 1. Kinetic parameters of HIV-1 and TEV PRs. A) His_6-MBP-VSQNY↓PIVQ-mTurquoise2 and His_6-MBP-VSQNY↓PIVQ-mApple substrates were used for activity measurements. Kinetic parameters were determined by non-linear regression of data. B) Kinetic parameters determined for the HIV-1 PR using an oligopeptide substrate.

A)

	Fluorescent protein	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} s^{-1})
HIV-1 PR	mTurquoise2	0.321±0.001	0.0163 ± 0.0001	19.69 ± 0.14
	mApple	0.050± 0.0008	0.0027 ± 0.0002	18.52 ± 1.40
TEV-PR	mTurquoise2	0.0050 ± 0.0001	0.0024 ± 0.0003	2.08 ± 0.26
	mApple	0.0038 ± 0.0001	0.0021 ± 0.0004	1.81 ± 0.35

B)

	oligopeptide	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} s^{-1})
HIV-1 PR	VSQNY↓PIVQ	32.21 ± 5.34	2.02 ± 0.51	15.94 ± 4.82